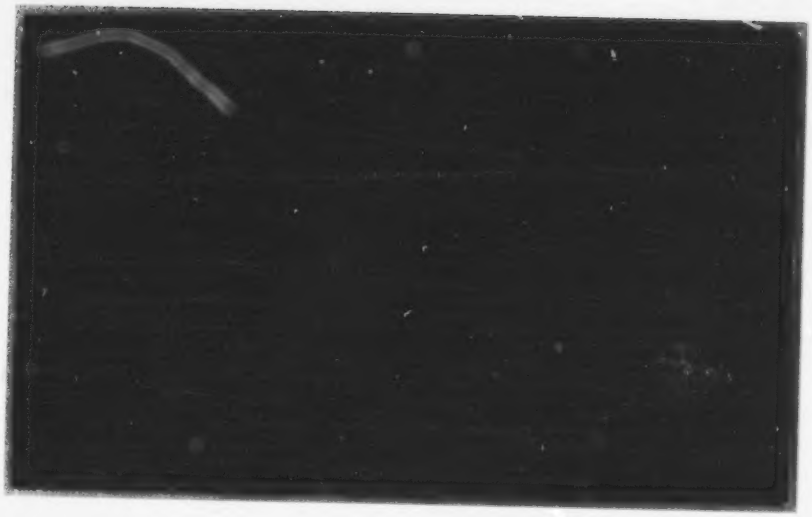
CENTRE DE RECHERCHE SUR LES TRANSPORTS
(C.R.T.)
CENTRE FOR RESEARCH ON TRANSPORTATION

UNIVERSITÉ DE MONTRÉAL
ÉCOLE DES HAUTES ÉTUDES COMMERCIALES
ÉCOLE POLYTECHNIQUE



CENTRE DE RECHERCHE SUR LES TRANSPORTS
UNIVERSITÉ DE MONTRÉAL
C.P. 6128, SUCCURSALE A,
MONTRÉAL
CANADA H3C 3J7

TÉLÉPHONE: (514) 343-7575
TÉLEX: (21) 05-561325
EMAIL: crt@crt.umontreal.ca
TÉLÉCOPIEUR (FAX): (514) 343-7121

CRT-734

From Correlation to Distributed Contiguities: A Family of AR-C-D Autocorrelation Processes

by

Ulrich C.H. Blum*
Denis Bolduc**
and
Marc J.I. Gaudry***

* Universität Bamberg, Volkswirtschaftslehre, insbesondere Wirtschaftspolitik, Postfach 1549, 8600 Bamberg, Germany.
** Département d'économique, Université Laval.
***Département de sciences économiques et Centre de recherche sur les transports, Université de Montréal.

The research for this paper was supported in part by the Ministère de l'enseignement supérieur, de la science et de la technologie du Québec and by the Natural Sciences and Engineering Research Council of Canada. Ulrich Blum and Marc Gaudry also benefited from the Alexander von Humboldt Foundation programs. Comments by Marcel G. Dagenais and Leen Hordijk were also very useful.

Universität Bamberg
Volkswirtschaftliche Diskussionsbeiträge — Nr. 38
Université Laval
Département d'économique — Cahier #9019
Université de Montréal
Centre de recherche sur les transports — Publication #734
Département de sciences économiques — Cahier #9031

November 1990

91-01372

Dépôt légal – 4e trimestre 1990

ABSTRACT

We distinguish between correlation and autocorrelation among values of regression variables or residual errors and define a family of autoregressive, contiguous and distributed processes to extract the systematic information that they contain using discrete Boolean impact matrices. We focus on two cases, called the R-Koyck and R-Almon first-order specifications. The former is shown to be equivalent to a constrained R-ARMA process and provides an alternative interpretation of the MA part. It is illustrated with a first order, infinite degree, $G(\pi)$-proximate $(1,\infty,G(\pi))$ irregular spatial example. Applicability of the family to other irregular processes with natural and directed orderings is discussed.

Keywords: RESIDUAL ERRORS, CONTIGUITY, SPATIAL AUTOCORRELATION, PROXIMITY, DATA ORDERING, IRREGULAR STOCHASTIC PROCESSES, ALMON, KOYCK, ARMA PROCESS INTERPRETATION.

RÉSUMÉ

Nous distinguons entre corrélation et autocorrélation entre les valeurs des variables ou des erreurs résiduelles d'une régression et formulons une famille de processus autoregressifs contigus et distribués afin d'en extraire, grâce à des matrices Booléennes d'impact, l'information systématique qui s'y trouve. Nous portons notre attention sur deux cas, nommés R-Koyck et R-Almon d'ordre un, et montrons que le premier est équivalent à un processus R-ARMA contraint dont il fournit une autre interprétation de la composante MA. Nous l'illustrons à l'aide d'un exemple spatial irrégulier du premier ordre, de degré infini et π-proche $(1,\infty,G(\pi))$. Nous discutons l'application de cette famille à d'autres données ordonnées naturellement ou directement par le chercheur.

Mots-clés: ERREURS RÉSIDUELLES, CONTIGUITÉ, AUTOCORRÉLATION SPATIALE, PROXIMITÉ, DONNÉES ORDONNÉES, PROCESSUS ALÉATOIRES IRRÉGULIERS, ALMON, KOYCK, INTERPRÉTATION DES PROCESSUS ARMA.

ZUSAMMENFASSUNG

Wir unterscheiden zwischen Korrelation und Autokorrelation unter den Werten von Regressionsvariablen oder Fehlerresiduen und definieren eine Familie von autoregressiven, benachbarten und verteilten Prozessen, um mit Hilfe diskreter Bool'scher Einwirkmatrizen die systematische Information, die diese enthalten, auszuwerten. Wir lenken dabei unsere Aufmerksamkeit auf zwei Fälle, die wir als R-Koyck- und als R-Almon-Spezifikationen erster Ordnung bezeichnen. Erstere ist äquivalent zu einem beschränkten R-ARMA-Prozeß und erlaubt eine alternative Interpretation des MA-Teils. Sie wird mit Hilfe eines irregulären räumlichen Beispiels $(1,\infty,G(\pi))$ erklärt, das durch Nachbarschaft erster Ordnung, unendlichen Grad der Einwirkung und eine Gewichtungsfunktion $G(\pi)$ charakterisiert ist. Die Möglichkeit der Anwendung auf andere irreguläre Prozesse mit natürlicher oder dirigierter Ordnung wird diskutiert.

Schlüsselbegriffe: Residualfehler, Nachbarschaft, räumliche Autokorrelation, Ordnung der Daten, irregulärer stochastischer Prozeß, Almon, Koyck, Interpretation eines ARMA-Prozesses

CONTENTS

1. INTRODUCTION 1

2. FRAMEWORK: IMPACT MATRICES 1

3. THE AR-C-D FORMULATION 3

A. From correlation to autocorrelation 4
B. Natural or directed order 5
C. Contiguity degree of the order *l* 6
D. Distributed proximity of the contiguity sequence 6

4. CASE 1: THE FIRST ORDER R-KOYCK FORMULATION 7

A. Preliminaries: normalization of *R* impact matrices 7
B. Specification 8
C. Properties 9
D. R-Koyck or R-ARMA(1,1)? 11
E. Numerical example 12

5. CASE 2: THE FIRST ORDER R-ALMON FORMULATION 16

6. EXTENSIONS AND OTHER CONSIDERATIONS 17

A. Naturally and directly ordered data 18
B. Residual errors and/or variables 19

7. CONCLUSION 20

APPENDIX 1. Relationships between $\underline{R}$ and $\underline{R}^*$ matrices 21

APPENDIX 2. Data and *R* impact matrix for Munich, 1980 23

REFERENCES. 24

- - - - - - - - -

TABLE 1. Specification of an AR-C-D process 3

TABLE 2. Cobb-Douglas production function estimates, Munich 1980 15

1. INTRODUCTION

In this paper, we define a family of **autoregressive**, **contiguous** and **distributed** processes (AR-C-D) to analyze the dependencies among the realizations of a variable and similarly among the residual errors in a regression. These AR-C-D processes make extensive use of weighted power series of Boolean impact matrices to extract systematic information on these dependencies. Such a use of the Boolean matrices may be seen as an alternative to specifications of impact matrices made in terms of continuous variables and to specifications of multiple-order autoregressive processes. Among a wide range of potential candidates, we focus on two cases, namely the R-Koyck and the R-Almon first-order specifications. The former, which we denote as the AR-C-D$(1,\infty,G(\pi))$ process, is a first-order autoregressive process characterized by a contiguity polynomial of infinite degree with weights distributed according to $G(\pi)$. As we will see, this process will turn out to be observationally equivalent to a constrained R-ARMA(1,1) process. This AR-C-D process provides an alternative interpretation of the moving average component of standard ARMA models and of its relationship to the inverted autoregressive structure interpretation. After describing the properties of this first process, we provide a numerical illustration where we estimate a production function based on cross-sectional data with errors assumed to be generated from an AR-C-D$(1,\infty,G(\pi))$ process. We then briefly discuss the applicability of the process to other problems such as time series models and to models in other fields than economics, as well as its possible use on specific variables in regression models. The second member considered has an Almon structure which we define as the AR-C-D$(1,s,A(L))$ process.

2. FRAMEWORK: IMPACT MATRICES

To situate the problem, consider the standard linear regression model to which we add a simple autoregressive error process. We write it as

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots + \beta_K X_K + v, \qquad (1)$$

$$v = \rho R v + w, \tag{2}$$

where $y, X_k, k = 1,\ldots,K$ are $(N \times 1)$ vectors of observations, v and w are $(N \times 1)$ vectors of errors, $\beta_k, k = 1,\ldots,K$, and the correlation coefficient ρ are parameters and R is a $(N \times N)$ impact matrix, also referred to as a weighting matrix in the spatial correlation literature. The element r_{it} in the autoregressive process detailed to the level of a single observation,

$$v_t = \rho \sum_{i=1}^{N} r_{it} v_i + w_t, \quad t = 1,\ldots,N, \tag{3}$$

is defined as

$$r_{it} = \begin{cases} 0, \text{ by convention for } t = i, \\ \text{value established by an impact criterion, } t \neq i. \end{cases} \tag{4}$$

For simplicity, we assume that the sample contains no missing observation. Clearly, the r_{it} elements of the impact matrix may be used to describe the influence of each error upon the others. The nature of this influence may be due to missing regressors, improper specification of the functional form (e.g. Gaudry and Wills, 1978) or of the error variance (Dagenais et al., 1987), or other reasons usually linked to mispecification. Practitioners suspect that this R matrix is in some sense "full" and know that the extraction of its systematic information has great importance even if the finite samples properties of the resulting estimators are unclear, as they are bound to belong to models that are incorrectly specified. Here we are more concerned with developing a heuristic approach to modeling the residual error - that we assume to result from specification errors - than we are concerned with searching for the "true" specification.

3. THE AR-C-D FORMULATION

A specific AR-C-D process can be put forward after answering the four different questions displayed in Table 1.

TABLE 1

Specification of an AR-C-D process

Question asked	Realization	Value
A. What equidistance rule defines auto-correlation?	Order, defined by R_l, a Boolean matrix for order l.	As many orders as desired: $l = 1,\ldots,L$.
B. Do natural (time, space, biological status, etc.) or socioeconomic factors define equidistance?	Meaning of the order l.	Natural or directed order (NO or DO).
C. Does realization in t depend on equidistant neighboring realizations only, or also on less contiguous ones (in a chained fashion)?	Degree of the polynomial series on impact matrix R_l, as defined by the sequence: $R_l, R_l^1, \ldots, R_l^{s_l}$.	Length of contiguity tail as desired: $s_l \in \{1,\ldots,\infty\}$ for impact matrix R_l associated with order l.
D. Are the degrees of proximity weighted in some distributed fashion?	A distribution $G(\cdot)$ for the proximity measure π_l.	To be determined by the modeler.

Assume for example that we are interested in estimating the parameters of a production function where the error term is postulated to be spatially correlated. We shall approach the problem by answering these four questions in turn.

A. From correlation to autocorrelation

A standard approach in modeling the error covariance structure is to formulate an autoregressive process as in Equation (2) with an R impact criterion (RIC) **function of continuous variables** (Bolduc et al., 1989 and Bolduc et al., 1990). More precisely, to model the correlation, one can always use a complex function of distances or a simpler function such as

$$r_{t.} = \left(\frac{1}{d_{t,1}^{\gamma}} \quad \cdots \quad \frac{1}{d_{t,t-1}^{\gamma}} \quad 0 \quad \frac{1}{d_{t,t+1}^{\gamma}} \quad \cdots \quad \frac{1}{d_{t,N}^{\gamma}} \right), \tag{5}$$

for the t^{th} row, where $\gamma > 0$ is a parameter to be estimated. In such a case, the parameter ρ can be interpreted as a spatial correlation parameter associated with a distance criterion function. The alternative approach that we choose to follow here is to approximate this process based on continuous proximity measures by a weighted sum of increasing powers of the Boolean matrix associated with a given order of contiguity. As noted in Anselin (1988), a Boolean impact matrix can always be interpreted as a portrayal of the structure of linkages in a network. The powered matrix shows on its off-diagonal the total number of connecting paths, of length equal to the power, between each pair of nodes in the network. We now adopt the following terminology.

An order l, $l = 1, \ldots, L$, is established by applying a specific equidistance rule as criterion in (4). This yields an impact matrix that we denote R_l. The equidistance rule may be a function of time, space, or other dimensions defined by given variables, for instance

Let r_{ti} = 1 if zones t and i are adjacent, and = 0 otherwise.	(NORIC-1)

Usually the first order process is associated with the shortest distance defined by the rule. The following two rules might well be used to define **two distinct contiguity orders**:

Let $r_{1,ii} = 1$ if zones t and i can be reached by crossing at most 3 borders using a shortest path algorithm, and $= 0$ otherwise.	(NORIC-2A)

Let $r_{2,ii} = 1$ if zones t and i can be reached by crossing at least 3 and at most 5 borders using a shortest path algorithm, and $= 0$ otherwise.	(NORIC-2B)

Using this approach, the first question in Table 1 can be answered by applying the assignment rule A_l that maps an equidistance function $E(\cdot)$ into the Boolean R_l matrix.

$$[r_{i,ii} = 0 \text{ or } 1] \quad <\!\!\!

and 2B are borderline cases of natural orders because a shortest path is part of the criterion function (6). The second question to ask is therefore whether natural (e.g. time, space, biological status) or directed (e.g. socioeconomic variables such as income, party affiliation, religion) factors define equidistance. Mixed orders are obviously possible.

C. Contiguity degree of the order l

The next question we ask is whether the realization of the error in zone t not only depends on equidistant "neighboring" realizations specified by the matrix R_l, but also on higher degree neighbors. To model this, we use a sequence of contiguity matrices defined in a chained fashion:

$$\{R_l^c\}_{c=1}^{s_l} = \left(R_l, R_l^2, R_l^3, \ldots, R_l^c, \ldots, R_l^{s_l}\right), \tag{7}$$

where c is assumed to be a nonnegative integer and s_l denotes the maximum degree allowed for contiguity of order l. As indicated in section A, R_l^2 captures the second degree neighbors (neighbors of neighbors) defined by the impact matrix R_l associated with the first neighbors of order l. In a spatial model, contiguity of degree c means that circularity and redundancy effects are possible in the sense that, if under NORIC-1 region 1 influences region 2 and the latter influences regions 1 and 3, then region 3 will influence regions 1, 2 and other regions: forward and backward linkages are allowed. In the first example examined below, s_l is $+\infty$ and, in the second case, it denotes a finite value.

D. Distributed proximity of the contiguity sequence

What is the relative importance of each component in the contiguity sequence (7)? In the first case considered, we assume that the dependency on higher degrees of contiguity decreases geometrically according to the normalized series of non-negative weights:

$$a_{lc} = a_l^{c-1}(1-a_l), \quad c = 1, \ldots, \infty \text{ and } a_l \in (0,1). \tag{8}$$

The normalization ensures that

$$\sum_{c=1}^{\infty} a_{ic} = \sum_{c=1}^{\infty} a_i^{c-1}(1-a_i) = \frac{1-a_i}{1-a_i} = 1. \tag{9}$$

To obtain a more intuitive measure that increases rather than decreases with the level of proximity, we define the following transformation:

$$\pi_i = 1 - a_i, \quad 0 \le \pi_i \le 1. \tag{10}$$

The reader will notice that the proposed structure is a straightforward extension of the usual Koyck distributed lag structure well-known in time-series to the matrix case. The second process that we consider later in the paper extends to the matrix case the traditional scalar Almon lag scheme.

4. CASE 1: THE FIRST ORDER R-KOYCK FORMULATION

A. Preliminaries: normalization of R impact matrices

Now assume that

$$\underline{R} \equiv \text{row normalized matrix } R, \tag{11}$$

i.e. each line of R is divided by its sum. In matrix form, $\underline{R}$ can be computed as $\underline{R} = R \cdot B^{-1}$ where B is a diagonal matrix with element b_{ii} being equal to the sum of the i^{th} row of R. The first purpose of this operation is to impose that the impact matrix raised to an increasing power eventually converges to a finite matrix. The process is therefore characterized as being **stable** as long as ρ belongs to the open unit interval ($\rho \in]-1, 1[$). The second purpose is to eventually guarantee a convex likelihood function over the stable unit interval of ρ, as Bolduc (1987) has shown.

Other types of normalization have been considered in the past for the contiguity matrices. Although it is not a requirement, the R matrix is very often symmetric. When the row normalization is applied to a symmetric R matrix, the resulting matrix is no longer symmetric. For models with large number of observations, to retain the property of symmetry can be computationally very efficient. Bolduc (1987) has therefore suggested to use $B^{-1/2}RB^{-1/2}$ as the normalized contiguity matrix. The advantage is that it is symmetric and possesses the same eigenvectors and eigenvalues as $\underline{R}$. The stability conditions of the symmetric normalized process are thus identical to the stability conditions of the process associated with $\underline{R}$. Since we do not force R to be symmetric in the present study, we favor the row normalized structure.

As for economic reasons the original Boolean matrix is not full, both of these normalization procedures create an implicit weighting system that depends on the number of elements per row or column. As an alternative, it may be useful to normalize by dividing each element of the Boolean matrix by the maximum number of neighbors across all observations. This also creates another non-neutral weighting system which is stable if ρ remains in the open unit interval.

B. Specification

Using the natural criterion NORIC-1, contiguity sequence (7) with $s \to \infty$ and distributed weights (8), after dropping the $l = 1$ subscript for convenience, we replace the autoregressive process in Equation (2), with the following stable autoregressive contiguous geometrically distributed process of order 1, degree ∞ and measure $G(\pi)$:

$$v = \rho(1-a)\sum_{c=1}^{\infty} a^{c-1}\underline{R}^{c}v + w, \tag{12}$$

$$= \rho(1-a)[\underline{R} + a\underline{R}^2 + a^2\underline{R}^3 + \ldots]v + w, \tag{13}$$

$$= \rho(1-a)[I + a\underline{R} + a^2\underline{R}^2 + \ldots]\underline{R}v + w, \tag{14}$$

$$= \rho(1-a)[I - a\underline{R}]^{-1}\underline{R}v + w, \tag{15}$$

or, from Equation (10), as

$$v = \rho \underline{R}^{*} v + w, \tag{16}$$

where $\underline{R}^{*}$ is defined as $\underline{R}^{*} = \pi[I - (1-\pi)\underline{R}]^{-1}\underline{R}$.

C. Properties

What are the properties of our AR-C-D $(1, \infty, G(\pi))$ process? Firstly, it is clear that the addition of a single additional parameter, namely π, to Equation (2) allows for a very simple endogenization of the relative importance of near and distant effects: it will depend on the value of the proximity measure π. Secondly, limiting values of π are of particular significance. Before proceeding any further, note the following results that we derive in Appendix 1.

Let Λ be the diagonal matrix which contains on the diagonal the eigenvalues $\lambda_t, t = 1, \ldots, N$ of $\underline{R}$. Since $\underline{R}$ is a row normalized impact matrix, one can check that the maximal eigenvalue in absolute value ($|\lambda_{MAX}|$) is equal to +1 and is unique. We arbitrarily choose this eigenvalue to be λ_1. Let also C denote the matrix of the eigenvectors associated with the eigenvalues λ_t. We now assume that $\underline{R}$ can be diagonalized[1]. We therefore can write

$$\underline{R} = C \Lambda C^{-1}.$$

In Appendix 1, it is proven that $\underline{R}^{*}$ and $\underline{R}$ share the same eigenvectors. It is also proven that the eigenvalues γ_t that diagonalize $\underline{R}^{*}$ are closely related to the λ_t eigenvalues. In fact, we show that

$$\underline{R}^{*} = C \Gamma C^{-1},$$

with Γ being the diagonal matrix that contains the γ_t eigenvalues, where :

---

1 See Appendix 1 for details.

$$\gamma_1 \equiv \lambda_1 \equiv 1 \text{ and } \gamma_t = \frac{\pi\lambda_t}{1-(1-\pi)\lambda_t}, \; t = 2,\ldots,N. \tag{17}$$

Limiting behavior when $\pi \to 1$

If the proximity measure π is at its highest value ($\pi \to 1$), it immediately follows from Equation (17) that the process in Equation (16) collapses to[2]

$$v = \rho \underline{R} v + w, \tag{18}$$

which means that large values of π (or, equivalently, small values of a) increase the impact of near neighbors to the point that only adjacent ones matter. In this case, the elements of the $\underline{R}^*$ matrix are simply equal to: 1/(number of non-zero elements in the row considered).

Limiting behavior when $\pi \to 0$

By contrast, as the proximity measure reaches its lowest value ($\pi \to 0$), or equivalently as a goes to 1, a greater relative weight is given to the tail. Using Equation (17), one directly obtains:

$$\lim_{\pi \to 0} \underline{R}^* = \lim_{\pi \to 0} C\Gamma C^{-1} = CE_{11}C^{-1} \equiv \overline{\underline{R}},$$

2 This is the case considered by Ord (1975) and Cliff and Ord (1981).

where E_{11} is the elementary matrix with a one in position 1,1 and zero elsewhere. Noting that the eigenvector associated with $\lambda_1 \equiv 1$ is the N-dimensional unit vector and calling c^1 the first row of C^{-1}, one can readily check that

$$\underline{\bar{R}} = \begin{bmatrix} c^1 \\ c^1 \\ . \\ c^1 \end{bmatrix}, \tag{19}$$

a matrix with identical rows. When $\pi \to 0$, we can therefore rewrite Equation (16) as

$$v = \rho \underline{\bar{R}} v + w, \tag{20}$$

with $\underline{\bar{R}}$ defined in Equation (19).

These two limiting cases yield the following bounds for the elements of $\underline{R}^* \equiv \pi[I - (1-\pi)\underline{R}]^{-1}\underline{R}$, used in Equation (16):

$$\underline{r}^*_{ti} \in (\underline{r}_{ti}, \underline{\bar{r}}_{ti}), \quad t, i = 1, 2, \ldots, N. \tag{21}$$

This relationship is very useful because one can then compute the maximum contribution of each error in explaining a corresponding error for any contiguity structure depicted by a given R impact matrix.

D. R-Koyck or R-ARMA(1,1)?

Rewrite (16) as follows:

$$[I - (1-\pi)\underline{R}]v = \rho\pi\underline{R}v + [I - (1-\pi)\underline{R}]w \tag{22}$$

or

$$[I - ((1-\pi) - \rho\pi)\underline{R}]v \quad = \quad [I - (1-\pi)\underline{R}]w. \tag{23}$$

This is of the form

$$[I - \varphi\underline{R}]v \quad = \quad [I - \theta\underline{R}]w, \tag{24}$$

with

$$\varphi = \theta - \rho\pi, \text{ and } \theta = 1 - \pi. \tag{25}$$

This clearly defines a restricted R-ARMA(1,1) formulation. Our R-Koyck specification can be seen as a handy parameterization of the ARMA(1,1), because π and ρ are very meaningful. From (25), one can obtain either φ and θ or ρ and π. Note however that the two models are observationally equivalent.

Macroeconomists emphasize that the MA part of these models represents random shocks (innovations) and the AR part denotes a "systematic" part. One might prefer to interpret that MA as the result of a geometrically decreasing AR part. This implies that innovations can be viewed as generating the growing AR tail from $-\infty$ to the present.

E. Numerical example

An insufficiently specified model

To illustrate our approach, imagine a regional labor market with several city-regions at the center and several rural regions at the fringe. Data are available on the number of employees and on the stock of capital in each region, but not on the number of those commuting from the fringe to the city core.

If we estimate a production function using these data, we will overestimate the contribution of labor in the city and underestimate it at the periphery: our production function could even exhibit increasing returns to scale, if the core is relatively large and is a net importer of labor.

We estimate such a function from data for eleven counties and cities which form the Munich labor market. Because of strong a priori knowledge on the functional form for Germany as a whole (Blum, 1986) we use a Cobb-Douglas function. As indicated in Appendix 2, the first and second regions are the city and the county of Munich, and regions three to eleven form a ring around this core.

But this is clearly not a well-specified model. We expect the residual errors of the regions to be correlated. To correct for this, our impact matrix R shown in Appendix 2 combines simple contiguity with hierarchy, as it contains a 1 only if the region within each of the groups is adjacent to a region of the same group, and a 0 otherwise:

Each residual error is correlated with its adjacent neighbors within the same level of spatial hierarchy.	(NODORIC-1)

This residual error impact criterion is a mixture of natural ordering (NO), given by spatial adjacency, and direct ordering (DO) resulting from a hierarchical rule - hence, its mixed name! The spatial example given here pertains to geographical interactions, but the process considered has wider potential applicability: for instance to product differentiation.

The likelihood function

Under the assumption that error w in Equation (16) is a normally distributed white noise, the logarithm of the likelihood function L, concentrated with respect to the coefficients β and the variance σ^2, is

$$\ln(L) = -\frac{N}{2}[1 + \ln(2\pi)] - \frac{N}{2}\ln(\hat{\sigma}(\rho,\pi)) + \ln|\det\{I - \rho\underline{R}^*\}|, \tag{26}$$

where an equivalent expression for the last term is[3]

$$\sum_{i=1}^{N} \ln|[1-\rho\gamma_i]| \;=\; \sum_{i=1}^{N} \ln[1-\rho\gamma_i], \tag{27}$$

if we use γ_i, the eigenvalues of the impact matrix $\underline{R}^* = \pi[I - (1-\pi)\underline{R}]^{-1}\underline{R}$, and the fact that $\rho\gamma_i < 1$, to compute the required determinant. The use of eigenvalues has two computational advantages: since the impact matrices are Boolean and depend only on the spatial setting, the eigenvalues can be computed once and for all prior to estimation; for large data sets, one can approximate Equation (27) by evaluating the eigenvalues on a subsample with N_1 observations (denote it as $\sum_{i=1}^{N_1} \ln[1-\rho\gamma_i^*]$). Then expand the expression to the full sample as follows:

$$\sum_{i=1}^{N} \ln[1-\rho\gamma_i] \;\approx\; \frac{N}{N_1} \sum_{i=1}^{N_1} \ln[1-\rho\gamma_i^*]. \tag{28}$$

As in any irregular process, the covariances of the residual errors are not invariant with respect to location, stationarity needs to be redefined, a task performed by Bolduc (1987) who also examined the conditions under which maximum likelihood estimations under the normality assumption are consistent and asymptotically normally distributed. He showed for the case of $\pi = 1$ that, after row normalization, the likelihood function is continuous and has a unique

---

3 See Ord (1975) for details.

maximum for values of ρ between -1 and 1, as each point of indeterminacies ($\rho = 1/\gamma_i$) lies outside this range. Because of that, he could show that the resulting maximum likelihood estimators were consistent and asymptotically normal. We performed a careful grid search over the π dimension to observe that unicity of the log likelihood function held in the -1 to 1 range of ρ.

Estimation results

Table 2 contains three columns. The first column contains uncorrected estimates. The second column adds ρ according to NODORIC-1 and the third column presents the estimates of the full model where the proximity measure π is also estimated.

TABLE 2
Cobb-Douglas production function estimates,
Munich 1980

	(1)	(2)	(3)
Labor (a)	0.6756	0.4693	0.4177
Capital (b)	0.6317	0.5465	0.5393
Scale (a+b)	1.3073	1.0158	0.9570
Log-likelihood	-82.7175	-80.3634	-79.1788
R^2	0.9777	0.9912	0.9947
ρ	0.0000	0.7785	0.8686
$\pi = 1 - a$	1.0000	1.0000	0.0950

All estimates are significant at the 5% level

The most general model of Column 3 is accepted at the 5% level when it is compared to the uncorrected model of Column 1; however, the proximity measure π is surprisingly low, which implies an influence of relatively far regions, but does not provide much explanatory power over and above that already offered by ρ, which is positive and significant. It is also interesting to note that the uncorrected model implies increasing returns to scale, which is not very plausible, while the corrected models imply essentially constant returns to scale.

5. CASE 2: THE FIRST ORDER R-ALMON FORMULATION

It is in principle possible to consider lower degrees of contiguity than infinity and other distribution schemes than the geometric one used here, but it is not clear whether one would always obtain an expression as simple as in Equation (16). In this section we present a second AR-C-D process, namely the first order R-Almon formulation.

In economics, the Almon scheme is usually imposed on the lag structure of a given regressor in a regression. Here we use the Almon structure to model the lag structure of the residual errors of a regression. We define the first order R-Almon of degree L and measure $f(i)$ as:

$$v = \rho \sum_{c=1}^{L} \delta_c^* \underline{R}^c v + w, \tag{29}$$

$$= \sum_{c=1}^{L} \delta_c \underline{R}^c v + w, \tag{30}$$

$$= \delta_1 \underline{R}^1 v + \delta_2 \underline{R}^2 v + \ldots + \delta_L \underline{R}^L v + w, \tag{31}$$

where $\delta_c = \rho \delta_c^*$ is imposed for identifiability. The Almon lag scheme reduces parameterization by assuming that the δ's can be approximated by some function $\delta_c = f(c)$ representing a polynomial of suitable degree. For a third-degree polynomial, one would have:

$$f(c) = \alpha_0 + \alpha_1 c + \alpha_2 c^2 + \alpha_3 c^3,$$

which implies that

$$\begin{aligned}
\delta_1 &= \alpha_0 + \alpha_1 + \alpha_2 + \alpha_3 \\
\delta_2 &= \alpha_0 + 2\alpha_1 + 4\alpha_2 + 8\alpha_3 \\
\delta_2 &= \alpha_0 + 3\alpha_1 + 9\alpha_2 + 27\alpha_3 \\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\delta_L &= \alpha_0 + L\alpha_1 + L^2\alpha_2 + L^3\alpha_3
\end{aligned} \tag{32}$$

Substituting Equation (32) in Equation (31) and rearranging gives

$$\nu = \alpha_0\left(\sum_{c=1}^{L} \underline{R}^c\nu\right) + \alpha_1\left(\sum_{c=1}^{L} c\underline{R}^c\nu\right) + \alpha_2\left(\sum_{c=1}^{L} c^2\underline{R}^c\nu\right) + \alpha_3\left(\sum_{c=1}^{L} c^3\underline{R}^c\nu\right) + w. \tag{33}$$

Thus, with a third-degree polynomial, only four parameters need to be estimated to retrieve the $\delta_c, c = 1,\ldots,L$ parameters. More specific details on how to use and interpret the estimates of an Almon lag specification can be found in standard econometrics textbooks.

6. EXTENSIONS AND OTHER CONSIDERATIONS

The main idea underlying the analysis above involves the introduction of restrictions to specify the covariances of disturbances associated with a group of regions, with the correlation between any two regions declining geometrically, but more or less rapidly, according to an equidistance function determining who is a first-degree neighbor, and by implication who is a higher degree neighbor. This focus, useful for presentation purposes is clearly not unique: nonspatial interpretations and multivariate formulations are also of interest.

A. Naturally and directly ordered data

In the case of classical time series, adjacency in NORIC-1 is naturally given by assuming that the previous time period can influence the current one.

This means that, with one additional parameter π, one can take into account the influence of remote observations. In contrast with the procedure used with exponential smoothing, where the weights associated to past occurrences are fixed, the endogenous nature of π gives flexibility. In the case of $\pi = 1$, one obtains the classical first order spatial autoregressive (SAR(1)) specification. In time series, it might be possible to model behavior of the residual errors linked to a yearly peak: for instance, in an inventory model the residual errors for all months of the year might be related in decreasing fashion to the Christmas, or to the harvest time.

Assume that an inventory model of highly seasonal products is to be estimated, e.g. chocolate bunnies for Easter, chocolate Santa Clauses for Christmas, seasonal pralines, etc. One would expect an increasing built up of the respective stocks in the company inventory over the τ months preceding the occasion of the sales. A conventional way of treating this structure in time-series analysis would be to include autocorrelation of different orders, e.g. 12th order, which relates a value to that of the same month of the preceding year. We, however, propose

Let $r_{1,n} = 1$ if $i \in [t-\tau, t-1], t \geq \tau+1$, where t designates the month of sale of a seasonal product, and $= 0$ otherwise.	(NORIC-3)

Our distributed lag provides an estimate of the tail that corresponds to the different build ups of products over time.

Naturally ordered data often provide interesting structures. For instance, one could define, in a biological model, the first and second degrees of equidistance in terms of dominant and recursive characteristics or in terms of ancestors' traits, for instance:

Let $r_{1,st} = 1$ if i is the grand parent of t, and
$= 0$ otherwise. (NORIC-4A)

Let $r_{2,st} = 1$ if i is a parent of t suffering from sickle cell anemia, and
$= 0$ otherwise. (NORIC-4B)

With such cases, and with demographic data generally, the specification of degrees of contiguity is closer to natural language, as in "first degree cousins", etc.

Socioeconomic structures provide rich opportunities to specify equidistance functions (6) according to the researcher's direct intuition or modeling hypothesis that may be related to hierarchical or professional position, income or other available Z_D factors. For instance, Alessie and Kapteyn (1985) estimate an Almost Ideal Demand System where individuals place weight on the demand of a reference group of all other households.

B. Residual errors and/or variables

We have expressed ourselves as if equation (16) were only applicable to regression residual errors. However, we could have specified an unconstrained model in which each variable of the regression model (1) would have had its own AR-C-D structure. Such a multivariate formulation may be too complicated to implement and imposing the structure only to the residual errors can be seen as a more tractable and profitable[4] approximation to one that would include all variables, or to one that would be used on both variables and residual errors.

4 For instance, Case (1987) compares a demand model with spatial correlation applied to the residual errors to the same model with spatial correlation applied only to the dependent variable and finds little to choose between the two. She also finds that combining both

On the other hand, sometimes one is interested in examining Granger causality from X_t to y, a test that often implies testing the null hypothesis that all coefficients of the "lagged" values of X_t are jointly insignificant. Imposing an AR-C-D structure on X_t could provide a half-way house and a sturdier maintained hypothesis than that provided by a set of coefficients on a hard-to-define set of usually very collinear lagged regressors required to whitewash variables, a task often carried out outside of the regression model.

7. CONCLUSION

We have specified autoregressive processes with tails of varying lengths and shapes dependent upon a single-parameter proximity measure. We have suggested that combinations of such overlapping tails could be a useful alternative to full correlation matrices.[5] We have defined two AR-C-D processes and showed that the first one, illustrated numerically with an irregular first order spatial process defined on real data, suggested an alternative interpretation of MA processes as AR processes with geometrically declining tails. We have discussed other uses of the AR-C-D processes with other irregular processes defined by both natural and directed orderings.

---

processes gives no improvement. Finally, she has evidence of multiple solutions, which may be due to the apparent use of non normalized R matrices. It is also known that implicit restrictions imposed when adopting an AR(1) error term amount to assuming a largely identical temporal structure for all the variables involved in the specification and implies that these variables are mutually Granger non-causal (Spanos, 1987-88).

5 Note also that direct specification of the elements of the correlation matrix do not insure positive definiteness of the fitted correlation matrix. If a multiple order process such as ours is used, stability of each process does not imply stability of the combined processes.

APPENDIX 1.

Relationships between $\underline{R}$ and $\underline{R}^*$ matrices

In this appendix, we study the behavior of the $\underline{R}^*$ matrix defined as

$$\underline{R}^* = \pi[I-(1-\pi)\underline{R}]^{-1}\underline{R}. \qquad (A.1)$$

Recall that $\underline{R}$ designates the row normalized version of an original R impact matrix. In the spatial autocorrelation literature, the impact matrix R is usually assumed to be symmetric. As indicated in Bolduc (1987), the properties of stability on the unit interval of ρ, of consistency and asymptotic normality of the ML estimators remain unchanged when non symmetrical R are employed.

Below, the only condition imposed on R is that there exist a Z matrix that diagonalizes it. It is well known that any symmetric matrix can be diagonalized. However, a general square matrix cannot always be diagonalized. Under the condition that all the roots of R are distinct, a general $(N \times N)$ matrix can always be written as

$$R = Z\Lambda Z^{-1}, \qquad (A.2)$$

where Λ is a diagonal matrix which contains the eigenvalues $\lambda_1,\ldots,\lambda_N$. In what follows, we assume that (A.2) holds.

A first result used in Bolduc (1985, 1987) is that $\underline{R}$ and R share the same eigenvalues. This implies that for $\underline{R}$, we will be able to write

$$\underline{R} = C\Lambda C^{-1}, \qquad (A.3)$$

with an obvious notation. Again from Bolduc (1987), when we order the eigenvalues in decreasing order, we know the following relationship holds

$$\lambda_N \leq \ldots \leq \lambda_s \leq 0 \leq \lambda_{s+1} \leq \ldots \leq \lambda_2 < \lambda_1 \equiv 1, \qquad (A.4)$$

where λ_s denotes the largest negative eigenvalue and λ_{s+1} the smallest positive eigenvalue.

LEMMA 1

$\underline{R}^*$ and $\underline{R}$ share the same eigenvectors. Also, the eigenvalues γ_n, $n = 1,\ldots,N$ that diagonalize $\underline{R}^*$ are simple transformations of the eigenvalues λ_n, $n = 1,\ldots,N$ of the $\underline{R}$ matrix. Finally, as for $\underline{R}$, the largest eigenvalue in absolute value of $\underline{R}^*$ is $+1$.

proof

Using $(A.3)$, one can write

$$\begin{aligned}
\underline{R}^* &= \pi[I-(1-\pi)\underline{R}]^{-1}\underline{R} \\
&= \pi[I-(1-\pi)C\Lambda C^{-1}]^{-1}C\Lambda C^{-1} \\
&= \pi C[I-(1-\pi)\Lambda]^{-1}C^{-1}C\Lambda C^{-1} \\
&= \pi C[I-(1-\pi)\Lambda]^{-1}\Lambda C^{-1} \\
&= C\begin{bmatrix} \dfrac{\pi\lambda_1}{1-(1-\pi)\lambda_1} & 0 & \cdots & 0 \\ 0 & \dfrac{\pi\lambda_2}{1-(1-\pi)\lambda_2} & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & \dfrac{\pi\lambda_N}{1-(1-\pi)\lambda_N} \end{bmatrix} C^{-1}. \qquad (A.5)
\end{aligned}$$

Since $\lambda_1 \equiv 1$, we can finally write

$$\underline{R}^* = C\Gamma C^{-1}, \qquad (A.6)$$

with Γ being the diagonal matrix that contains the γ_n eigenvalues, where : $\gamma_1 \equiv \lambda_1 \equiv 1$, and $\gamma_n = \pi\lambda_n/[1-(1-\pi)\lambda_n]$, $n = 2,\ldots,N$. []

APPENDIX 2.

Data and *R* impact matrix for Munich, 1980

Nr.	Region	Gross value added (Mill. DM)	Population	Private capital assets (Mill. DM)
1	München-Stadt	53087	1298000	265949
2	München-Land	6385	245600	3145
3	Erding	1473	86200	1390
4	Ebersberg	1453	94900	1261
5	Freising	2144	112600	1563
6	Dachau	1546	101600	1346
7	Fürstenfeldbruck	2209	167700	1652
8	Starnberg	1904	106000	1677
9	Bad Tölz-Wolfratshausen	1818	97000	1761
10	Miesbach	1457	84400	1417
11	Rosenheim	3047	182000	3327

Adjacency matrix R

1	München-Stadt	0	1	0	0	0	0	0	0	0	0	0
2	München-Land	1	0	0	0	0	0	0	0	0	0	0
3	Erding	0	0	0	1	1	0	0	0	0	0	0
4	Ebersberg	0	0	1	0	0	0	0	0	0	0	1
5	Freising	0	0	1	0	0	1	0	0	0	0	0
6	Dachau	0	0	0	0	1	0	1	0	0	0	0
7	Fürstenfeldbruck	0	0	0	0	0	1	0	1	0	0	0
8	Starnberg	0	0	0	0	0	0	1	0	1	0	0
9	Bad Tölz-Wolfratshausen	0	0	0	0	0	0	0	1	0	1	0
10	Miesbach	0	0	0	0	0	0	0	0	1	0	1
11	Rosenheim	0	0	0	1	0	0	0	0	0	1	0

REFERENCES.

ALESSIE, R. and KAPTEYN, A. (1985) "Habit formation and Interdependent Preferences in the Almost Ideal Demand System", mimeo, Tilburg University, The Netherlands.

ANSELIN, L. (1988) Spatial Econometrics, Dordrecht, Kluwer Academic Publishers.

BLUM, U. (1986) "The Impact of Location and Centrality on Regional Income: The German Experience 1976-1980", in Space-Structure-Economy. A Tribute to August Lösch, R. Funck and A. Kuklinski, eds., v. Loeper Verlag, Karlsruhe, 219-228.

BOLDUC, D. (1987) "On the Estimation of Models with Generalized SAR(1) Processes on the Residuals of a Regression", publication # 508, Centre de recherche sur les transports, Université de Montréal.

BOLDUC, D., DAGENAIS, M.G. and GAUDRY, M.J.I. (1989) "Spatially Autocorrelated Errors in Origin-Destination Models: A New Specification Applied to Aggregate Mode Choice", Transportation Research B 23(5), 361-372.

BOLDUC, D., LAFERRIERE, R. and SANTAROSSA, G. (1990) "Autoregressive Error Components in Travel Flow Models", département d'économique, Université Laval.

CASE, A.C. (1987) "On the Use of Spatial Autoregressive Models in Demand Analysis", Discussion paper # 135, Research Program in Development Studies, Princeton University, October.

CLIFF, A.D. and ORD, J.K. (1981) Spatial Processes: Models and Applications, Pion Limited, London.

DAGENAIS, M.G., GAUDRY, M.J.I. and LIEM, T.C. (1987) "Urban Travel Demand: The Impact of Box-Cox Transformations with Nonspherical Residual Errors", Transportation Research B 21(6), 443-477.

GAUDRY, M.J.I. and BLUM, U.C.H. (1988) "An Example of Correlation Among Residuals in Directly Ordered Data", Economic Letters 26, 335-340.

GAUDRY, M.J.I. and WILLS, M.J. (1978) "Estimating the Functional Form of Travel Demand Models", Transportation Research 12(3), 195-211.

ORD, K. (1975) "Estimation Methods for Models of Spatial Interaction" Journal of the American Statistical Association 70(349), 120-126.

SPANOS, A. (1987-88) "Error Autocorrelation Revisited: The AR(1) Case", Econometric Reviews 6(2), 285-294.